



# Feature Profiles from Attribute Filtering for Classification of Remote Sensing Images

Minh-Tan Pham, Erchan Aptoula, Sébastien Lefèvre

## ► To cite this version:

Minh-Tan Pham, Erchan Aptoula, Sébastien Lefèvre. Feature Profiles from Attribute Filtering for Classification of Remote Sensing Images. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018, 11 (1), pp.249-256. 10.1109/JSTARS.2017.2773367 . hal-01672852

**HAL Id: hal-01672852**

**<https://hal.science/hal-01672852>**

Submitted on 13 Nov 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Feature profiles from attribute filtering for classification of remote sensing images

Minh-Tan Pham, *Member, IEEE*, Erchan Aptoula, Sébastien Lefèvre

**Abstract**—This paper proposes a novel extension of morphological attribute profiles (APs) for classification of remote sensing data. In standard AP-based approaches, an input image is characterized by a set of filtered images achieved from the sequential application of attribute filters based on the image tree representation. Hence, only pixel values (i.e. gray levels) are employed to form the output profiles. In this paper, during the attribute filtering process, instead of outputting the gray levels, we propose to extract both statistical and geometrical features from the connected components (w.r.t tree nodes) to build the so-called feature profiles (FPs). These features are expected to better characterize the object or region encoded by each connected component. They are then exploited to classify remote sensing images. To evaluate the effectiveness of the proposed approach, supervised classification using the random forest classifier is conducted on the panchromatic Reykjavik image as well as the hyperspectral Pavia University data. Experimental results show the FPs provide a competitive performance compared against standard APs and thus constitute a promising alternative.

**Index Terms**—Remote sensing imagery, attribute profiles, feature profiles, supervised classification, random forest

## I. INTRODUCTION

IMAGE classification is one of the most significant tasks in remote sensing imagery which serves for several applications in land use and land cover earth observation. Among a great number of proposed techniques in the literature (see a review in [1]), morphological attribute profiles (APs) [2] have been widely used thanks to their powerful multilevel modeling of spatial information from the image content and their efficient implementation via tree structures. By well preserving important spatial properties of regions and objects such as contours, shape, compactness, etc., APs characterize effectively the contextual information of the observed scene, hence relevant for remotely sensed image classification task.

So far, many studies have been contributed to exploit and extend the use of APs in this field [3]–[17]. For example, as the standard APs [2] were generated based on a min-tree and a max-tree, some of these studies have focused on advanced implementations using other tree structures such as the tree of shapes [4], [5] or the partition trees (binary partition tree,  $\alpha$ -tree,  $\omega$ -tree, etc.) [6]. Other works have proposed to extract the histogram [7] or the first-order statistical features [8], [9] of the local patch around each AP sample for better characterization of textures within high resolution images. Then, along with other efficient hyperspectral image feature

extraction techniques such as intrinsic image decomposition, manifold learning, recursive filtering, random walk collaborative learning, etc. [18]–[22], spatial-spectral approaches based on APs have drawn lot of attentions on researchers in recent years. Many studies have proposed to exploit AP-based spatial enhancement combined with spectral linear/non-linear feature extraction techniques such as the principal component analysis (PCA), kernel PCA, discriminant analysis feature extraction (DAFE), non-parametric weighted feature extraction (NWFE), etc. for classification of hyperspectral images [15]–[17]. For a recent survey, readers are referred to [23].

It should be noted that in all standard AP methods, the output profiles are a set of filtered images obtained by attribute filtering process. Hence, they are still the gray values of the connected components (CCs) w.r.t the nodes of the pruned tree (i.e. the tree whose nodes all verify the attribute filtering criteria). In other words, the various geometric and statistical properties of the CCs containing the pixels to be described, are captured only indirectly through the pixel values. Therefore in this paper, instead of reconstructing the filtered image (i.e. using pixel gray values) from the pruned tree, we propose to extract some statistical features (i.e. mean, standard deviation, entropy, etc.), together with some geometric and shape information (i.e. area, elongation, diagonal length of bounding box, etc.), of each CC to form the so-called feature profiles (FPs). These FPs are expected to better characterize the object or region enclosed by each CC than the classical APs. We note that the rest of this paper will focus on the generation of FPs based on min-tree and max-tree to compare with the standard APs. Similar approach can be undoubtedly applied to form FPs based on partition trees or tree of shapes.

The remainder of this paper is organized as follows. The next section recalls the background of APs. Section III overviews the extraction of different statistical and geometric features from tree nodes and then describes the construction of the proposed feature profiles. We present and discuss the results of our classification experiments in Section IV. Section V finally concludes the paper and addresses some future work.

## II. MORPHOLOGICAL ATTRIBUTE PROFILES

APs are multilevel image description tools obtained by successively applying a set of morphological attribute filters (AFs) [2]. Unlike usual image filtering operators which are directly performed on pixel level, AFs work on CC level based on the concept of image connectivity. In other words, an AF is a filtering operator applied on the tree's node level with regard to a specific attribute characterizing the size, shape, or other

M.T. Pham and S. Lefèvre are with the IRISA lab - Université de Bretagne Sud, UMR 6074, Vannes 56000, France. e-mail: minh-tan.pham@irisa.fr.

E. Aptoula is with the Institute of Information Technologies - Gebze Technical University, Kocaeli, 41400, Turkey

properties of objects present in the image. To this end, the generation of standard APs from an image can be summarized as a four-step process: 1) construct the max-tree (and the min-tree) from the image; 2) compute some relevant attributes (area, moment of inertia, standard deviation, etc.) from each node of the tree; 3) filter the tree by keeping/removing nodes according to their attribute values; 4) reconstruct the image from the filtered tree. Steps 3) and 4) can be done for different attributes (with different threshold values) to finally produce a set of filtered images (by stacking them) to form APs.

More formally, given a grayscale image  $X : E \rightarrow \mathbb{Z}, E \subseteq \mathbb{Z}^2$ , the standard generation of APs on  $X$  is achieved by applying a sequence of AFs based on a min-tree (attribute thickening operators  $\{\phi_k\}_{k=1}^K$ ) and a max-tree (i.e. attribute thinning operators  $\{\gamma_k\}_{k=1}^K$ ) as follows:

$$\text{AP}(X) = \left\{ \phi_K(X), \phi_{K-1}(X), \dots, \phi_1(X), X, \right. \\ \left. \gamma_1(X), \dots, \gamma_{K-1}(X), \gamma_K(X) \right\}, \quad (1)$$

where  $\phi_k(X)$  is the filtered image obtained by applying the attribute thickening  $\phi$  with regard to the threshold  $k$ . Similar explanation is made for  $\gamma_k(X)$ . As observed, the resulted  $\text{AP}(X)$  is a stack of  $(2K + 1)$  images including the original image,  $K$  filtered images from the thickening profiles and the other  $K$  from the thinning profiles. For more details about this AP computation, readers are referred to papers [2], [23].

Instead of calculating the APs based on both max-tree and min-tree image representation (i.e. original APs), other implementations have been proposed using the tree of shapes [4], [5] (i.e. self-dual APs) or the partition trees (such as  $\alpha$ -tree and  $\omega$ -tree to produce  $\alpha$ -profiles,  $\omega$ -profiles, respectively) [6]. In this work, we focus on the standard APs proposed in [2] using min-tree and max-tree. We note that the proposed strategy is applicable to all above-mentioned trees.

### III. FEATURE PROFILES FROM ATTRIBUTE FILTERING

In the above standard AP methods (as well as in all existing AP-based methods in the literature), the output profiles are a sequence of filtered images produced by attribute filtering process. They are still the gray values of the CCs w.r.t the nodes of the pruned tree (i.e. the tree in which all nodes verify the attribute filtering criteria). Equation (1) has defined the image profiles. Now, for each pixel  $p$  in the definition domain of  $X$ , we have:

$$\text{AP}(p) = \left\{ [\phi_K(X)](p), \dots, [\phi_1(X)](p), X(p), \right. \\ \left. [\gamma_1(X)](p), \dots, [\gamma_K(X)](p) \right\}. \quad (2)$$

From this equation, the vector  $\text{AP}(p)$  is formed by a concatenation of gray levels of filtered images. Even though it contains gray levels that result from (and depend on) a connected filtering sequence w.r.t. an arbitrary attribute, the resulting values provide only a limited “view”. A direct exploitation of this feature vector for classification task may lead to the lack of statistical properties of regions/objects related to the tree nodes. Therefore in this work, our motivation is to extract some statistical features (i.e. mean, standard deviation,

entropy, etc.), together with some geometric information (i.e. area, compactness, bounding box’s diagonal length, etc.), of each node, instead of outputting node’s gray level. We believe that these features can provide better characterization of the observed regions and objects, especially in remote sensing imagery. We now revisit some features or attributes which can be extracted from the tree nodes. Then, we describe the proposed method in details.

#### A. Node features

During the construction of APs, different attributes have been considered to model the spatial and structural properties of the object or region corresponding to a tree node. In particular, the following four attributes have been commonly exploited in most AP-based research studies:

- + *area* (which models the size of regions);
- + *moment of inertia* (which helps to discriminate elongated objects from compact ones);
- + *diagonal length of the bounding box* (which also models the region’s size);
- + *standard deviation* (which involves the region’s homogeneity).

Among them, the first three attributes involve the shape structures of the CC related to each tree node while the last one provides a statistical measure. In order to better characterize the node, many other features [24], [25] can be employed. They can be grouped into two main categories:

- + *statistical features*: related to the distribution of pixel values within the CC dedicated to the node. They can measure the smoothness, coarseness and regularity of the region surface. Some can be stated here such as mean, median, other first-order statistics (i.e. kurtosis, skewness, etc.), textures (i.e. entropy, correlation, homogeneity), etc.
- + *geometrical features*: related to the size and shape of the CC and not the pixel values within the CC such as elongation, perimeter, circularity, shape centroid, etc.

All of these features can be employed for attribute filtering process in any AP strategies. Here in our work, they are extracted to form the output profiles as well. The proposed strategy is now described.

#### B. Generation of FPs

As previously mentioned, in order to better characterize the region/object enclosed by the CC (i.e. which corresponds to a filtered tree’s node), node features are extracted instead of the node’s gray level. Fig. 1 provides an overview of how the proposed FP approach differs from the standard AP technique. In fact, after obtaining the pruned tree by an attribute filtering  $\phi_k$ , instead of reconstructing the filtered image using the nodes’ gray levels, we output different node features.

Specifically, for each pixel  $p$ , AP of  $p$  obtained by an arbitrary AF  $\phi_k$  is the gray value  $X'(p)$ , where  $X' = \phi_k(X)$  is the image reconstructed from the filtered tree (cf. Eq (2)). Now, let  $\Gamma_p(X)$  be the CC of  $X$  containing  $p$  and let  $m$  be a feature or an attribute, i.e. a function admitting a CC and outputting a real value, to be extracted. The FP of  $p$  will be  $m[\Gamma_p(X')]$ .

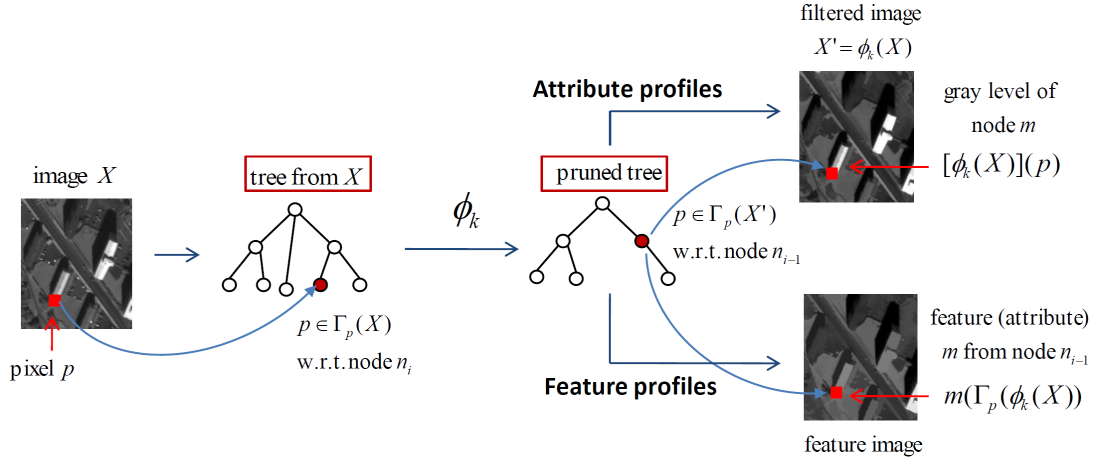


Fig. 1. Generation of the standard attribute profiles (APs) and the proposed feature profiles (FPs) from a grayscale image  $X$  w.r.t. the attribute filtering  $\phi_k$ . Here  $\Gamma_p$  denotes the connected component (CC) containing  $p$ ;  $X' = \phi_k(X)$  is the filtered image; and  $m$  is the feature (i.e. attribute) to be extracted.

More formally, the generation of FPs w.r.t the feature  $m$  based on min-tree and max-tree are defined as follows:

$$\text{FP}_m(p) = \left\{ m[\Gamma_p(\phi_K(X))], \dots, m[\Gamma_p(\phi_1(X))], X(p), m[\Gamma_p(\gamma_1(X))], \dots, m[\Gamma_p(\gamma_K(X))] \right\}. \quad (3)$$

Unlike in AP technique where only one profile is produced from one pruned tree, several features can be simultaneously extracted and stacked to provide the final FP:

$$\text{FP}_{m_1+\dots+m_n}(p) = [\text{FP}_{m_1}(p), \dots, \text{FP}_{m_n}(p)]. \quad (4)$$

Depending on a specific application, features should be chosen for relevancy. For example, geometrical shapes are more relevant for detection or tracking tasks. On the other hand, applications such as classification or recognition should be tackled by combining both statistical and geometrical information. For remote sensing classification task as considered in this work, we propose to exploit both of them. During the experimentation, we have evaluated the performance of different features separately as well as their combination. They will be reported in Section IV.

Last but not least, for classification of multi-channel remote sensing data, in particular hyperspectral images, the extended version of FPs (EFPs) is generated similarly to the extended APs (EAPs) [3]. In general, a feature extraction or feature selection method is first applied to reduce the dimensionality and remove redundant information from the image. For example, in the original work of EAPs [3], the authors proposed to first apply the PCA [26] and then compute APs from each of the first few components. Then, the final EAPs were constructed by stacking all the obtained APs. As mentioned in the introduction, other techniques have been proposed to replace the PCA for the feature extraction/dimensionality reduction task such as the independent component analysis (ICA) [27], kernel PCA (KPCA) [17], discriminant analysis feature extraction (DAFE) [15], non-parametric weighted feature extraction (NWFE) [28], etc. Here in our work, since the main purpose is to enhance the AP performance with FP

technique, the PCA is selected for standard EFP generation, but any adaptation or improvement using the above mentioned techniques can be undoubtedly applied.

#### IV. EXPERIMENTAL STUDY

This section describes our experimental study to evaluate the performance of the proposed method. Supervised classification has been carried out on both panchromatic and hyperspectral remotely sensed image data in order to confirm the effectiveness of FPs as well as the extended version (EFPs). We first introduce the two standard data sets which have been exploited in most studies related to APs and describe the experimental setup. Then, classification results will be provided with both qualitative and quantitative assessments.

##### A. Data sets

1) *Reykjavik data set*: The first data set is a panchromatic image of size  $628 \times 700$  pixels acquired by the IKONOS Earth imaging satellite with 1-m resolution in Reykjavik, Iceland. This data consists of six thematic classes including residential, soil, shadow, commercial, highway and road. The image was provided with already-split training and test sets (22741 training samples and 98726 test samples). The input image together with its thematic ground truth map for testing and training sets are shown in Fig. 2(a).

2) *Pavia University data set*: The second data set is the hyperspectral image acquired by the ROSIS airborne sensor with 1.3-m spatial resolution over the region of Pavia University, Italy. The image consists of  $610 \times 340$  pixels with 103 spectral bands (from 0.43 to 0.86  $\mu\text{m}$ ) and covers nine thematic classes: trees, asphalt, bitumen, gravel, metal sheets, shadows, meadows, self-blocking bricks and bare soil. For this image, 3921 training samples and 42776 test samples were split for classification task. The false-color image (made by combining the bands 31, 56 and 102), the ground truth map and the training set are shown in Fig. 2(b). As previously discussed, for this data set, we first performed the PCA on the image and the first four PCs (involving more than 99% of the total variance) were preserved for our experiments.

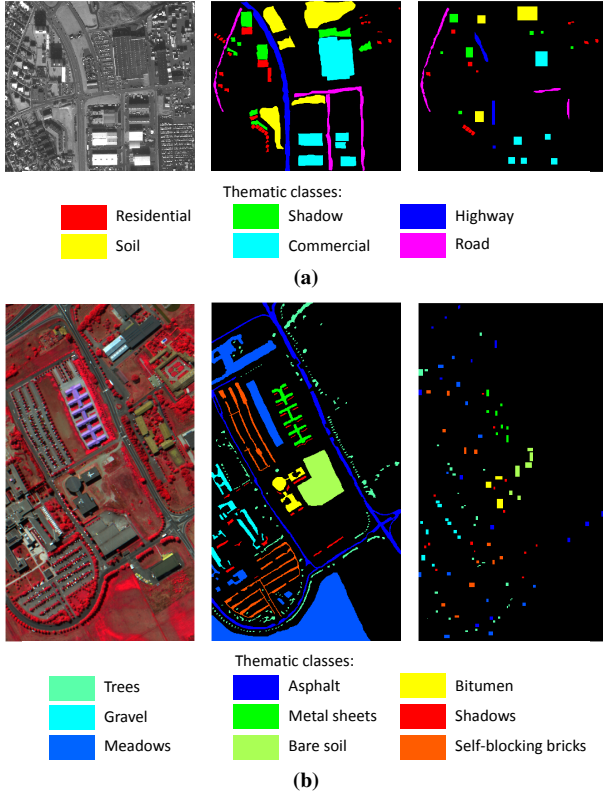


Fig. 2. Two data sets used in our experimental study. (a) The  $628 \times 700$  Reykjavik data (left to right: panchromatic, thematic ground truth with 6 classes and training set); (b) The  $610 \times 340$  Pavia University data (left to right: false-color image made by bands 31-56-102, ground truth including nine thematic classes and training set).

### B. Experimental setup

Supervised classification was conducted on our two data sets using the random forest classifier [29] with 100 trees. The number of variables involved in the training was set to the square root of the feature vector length. In order to evaluate and compare classification accuracy of different approaches, overall accuracy (OA), average accuracy (AA), and kappa coefficient ( $\kappa$ ) have been taken into account. The classification procedure was evaluated using the spectral information (PAN: panchromatic for Reykjavik and 4 PCs: four first PCA bands for Pavia University), the AP and FP features. For attribute filtering, we exploited two geometrical attributes including the area and the moment of inertia. Then for FP extraction, both statistical features (mean, standard deviation) and geometrical information (area) were employed and combined.

For area attribute, ten thresholds were adopted for the Reykjavik data as proposed by several papers [5], [13]. For the Pavia University data, they were automatically computed according to the work in [16]. We have:

$$\lambda_{a,Rey} = \{25, 100, 500, 1000, 5000, 10000, 20000, 50000, 100000, 150000\},$$

$$\lambda_{a,Pav} = \{770, 1538, 2307, 3076, 3846, 4615, 5384, 6153, 6923, 7692, 8461, 9230, 10000, 10769\}.$$

Next, the manual settings used in many studies [3], [14], [28]

were adopted for the moment of inertia attribute as follows:

$$\lambda_{i,Rey} = \lambda_{i,Pav} = \{0.2, 0.3, 0.4, 0.5\}.$$

### C. Classification results

In this section, experimental results are presented and discussed for each data set. For the Reykjavik data, some AP and FP output profiles are first displayed in Fig. 3 for visualization. Here, we show the profiles obtained by using three area thresholds  $\lambda_a = \{100, 5000, 50000\}$  for attribute filtering. AP technique (first row) outputted pixel's gray values while two FP techniques (second and third rows) extracted the area and the standard deviation features. As observed, FPs can provide some additional information that helps us to improve the classification performance reported in Table I. From the table, the best performance was produced by using both area and moment of inertia for attribute filtering and combining mean, standard deviation and area to generate the feature profiles ( $FP_{\mu+\sigma+a}$ ) with an OA equal to 86.23% ( $\kappa = 0.8246$ ). Compared to APs with area and moment attributes ( $AP_{a+i}$ ), an improvement of 3.69% in OA and 4.53% in  $\kappa$  was achieved. We remark that most FP techniques provide better performance than AP methods. For example, with area attribute filtering, using standard deviation as output feature ( $FP_{\sigma}$ ) can improve 3.2% of OA (84.73% compared to 81.53%) from the standard  $AP_a$  with the same feature length of 21. For another comparison, the best performance of the recent extinction profiles [13] for this specific Reykjavik data set was 85.58% in OA and  $\kappa = 0.8145$ , which is inferior than our best performance.

For better evaluation, Fig. 4 shows the classification maps and Table III provides the classification performance for each class (per-class accuracy) of Reykjavik data. Here, the results of AP and FP methods are compared in each case of using attribute filtering by area, moment and both area+moment. From the table, better results are provided by FPs for most of the classes (some FP results are lower but very few and slightly). In particular, a significant improvement (more than 10% for all three cases) is achieved for the soil class (i.e. yellow in the ground truth). From Fig. 4, this class was considerably mis-classified with the road class (in magenta) within three AP approaches. Readers can observe the soil region inside the white ellipse at the middle of the image. Meanwhile, this soil region was much better classified by FP approaches. This can be explained by the fact that statistical and geometrical features extracted from FPs are more discriminative than only gray levels employed by APs, especially in case that two classes have similar intensity levels.

Similar classification results are provided for the Pavia University image in Fig. 5, Tables II and IV. As for the previous data, the best classification performance was yielded by the extended FPs using mean, standard deviation and area as features ( $EFP_{\mu+\sigma+a}$ ) with OA= 97.01%, AA=98.34% and  $\kappa = 0.9607$ . Compared to its counterpart  $EAP_{a+i}$ , an improvement of 3.67% and 4.91% was achieved in terms of OA and kappa, respectively. Here, we can observe a similar behavior of AP and FP techniques to the Reykjavik data. The performance in terms of attribute filtering is: moment < area



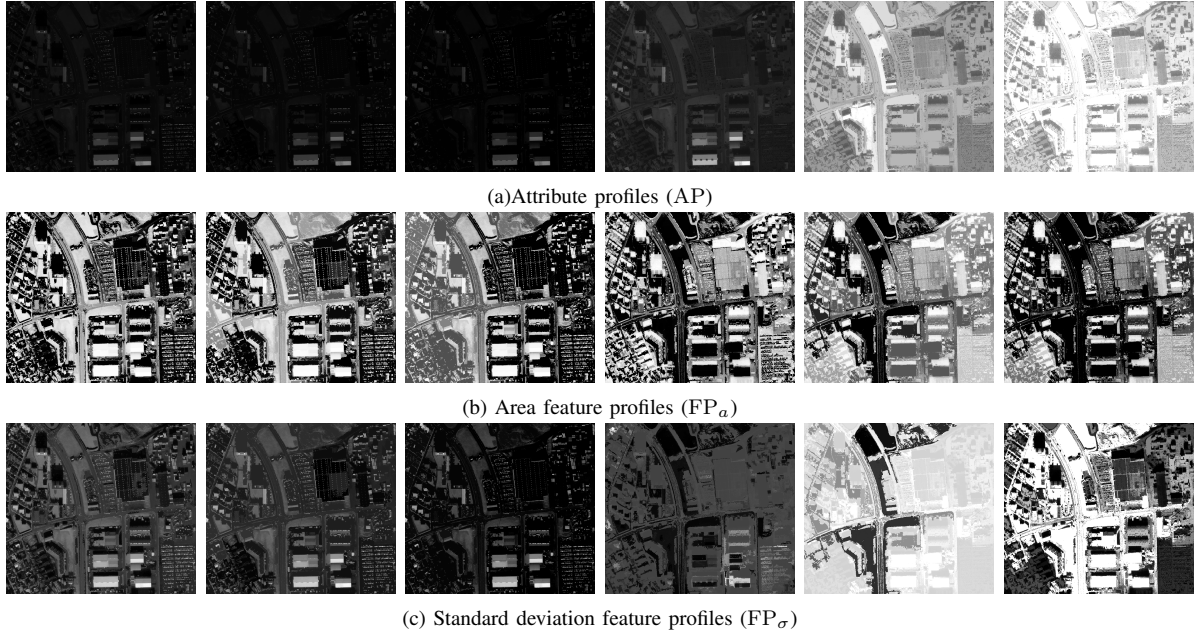


Fig. 3. Visualization of attribute profiles and feature profiles generated from the Reykjavik image using three area thresholds  $\lambda_a = \{100, 5000, 50000\}$ . The three left-side columns involve the closing (thickening) profiles and the three right-side ones involve the opening (thinning) profiles.

TABLE I  
CLASSIFICATION RESULT OF THE REYKJAVIK DATA OBTAINED BY DIFFERENT METHODS USING RANDOM FOREST WITH 100 TREES.

Method	Dimension	Classification result		
		OA (%)	AA (%)	$\kappa$
PAN	1	63.21	53.58	0.5237
<i>Attribute profiles</i>				
$AP_a$	21	81.53	77.41	0.7666
$AP_i$	9	72.51	68.50	0.6555
$AP_{a+i}$	30	<b>82.54</b>	<b>78.80</b>	<b>0.7793</b>
<i>Feature profiles with attribute 'area'</i>				
$FP_\mu$	21	80.12	77.66	0.7495
$FP_\sigma$	21	84.73	80.20	0.8056
$FP_a$	21	84.41	79.96	0.8021
$FP_{\mu+\sigma}$	42	85.03	80.53	0.8094
$FP_{\sigma+a}$	42	85.30	80.47	0.8130
$FP_{\mu+\sigma+a}$	63	<b>85.81</b>	<b>81.01</b>	<b>0.8191</b>
<i>Feature profiles with attribute 'moment of inertia'</i>				
$FP_\mu$	9	75.83	73.16	0.6962
$FP_\sigma$	9	75.31	72.60	0.6905
$FP_a$	9	74.11	71.76	0.6756
$FP_{\mu+\sigma}$	18	76.72	74.28	0.7076
$FP_{\sigma+a}$	18	79.42	75.80	0.7405
$FP_{\mu+\sigma+a}$	27	<b>79.98</b>	<b>76.29</b>	<b>0.7469</b>
<i>Feature profiles with 'area' + 'moment of inertia'</i>				
$FP_\mu$	30	80.92	78.45	0.7593
$FP_\sigma$	30	85.23	80.65	0.8113
$FP_a$	30	84.60	80.35	0.8045
$FP_{\mu+\sigma}$	60	84.73	80.52	0.8060
$FP_{\sigma+a}$	60	86.11	81.32	0.8231
$FP_{\mu+\sigma+a}$	90	<b>86.23</b>	<b>81.25</b>	<b>0.8246</b>

< area+moment. Then, for output features, a combination of  $\mu + \sigma + a$  has produced the best performance. However, a combination of two features such as  $\mu + \sigma$  or  $\sigma + a$  could also provide good results. As observed from Table IV, per-class accuracy was improved for most classes using FPs instead of

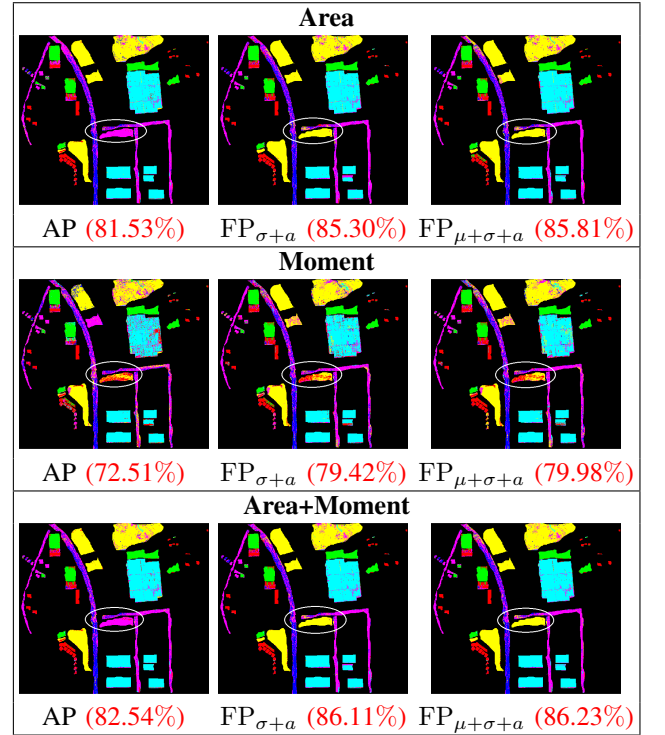


Fig. 4. Classification maps of Reykjavik data obtained by different AP and FP configurations. Each red number is OA produced by the related method.

the standard APs. In particular, by using both area+moment, FPs can enhance 9.95% of gravel class and more impressively, 25.15% for bare soil class. This issue again emphasizes the effectiveness of statistical and geometrical features provided by the proposed FP strategy.

Lastly, to enrich the discussion on this hyperspectral data, some experiments were also conducted in order to confirm the superior performance of FP against AP approach in case

TABLE II  
CLASSIFICATION RESULT OF THE PAVIA UNIVERSITY DATA OBTAINED BY  
DIFFERENT METHODS USING RANDOM FOREST WITH 100 TREES.

Method	Dimension	Classification result		
		OA (%)	AA (%)	$\kappa$
4 PCs	4	69.32	79.21	0.6226
<i>Attribute profiles</i>				
EAP <sub>a</sub>	116	93.09	94.30	0.9084
EAP <sub>i</sub>	36	78.83	88.04	0.7322
EAP <sub>a+i</sub>	152	<b>93.34</b>	<b>94.39</b>	<b>0.9116</b>
<i>Feature profiles with attribute 'area'</i>				
EFP <sub><math>\mu</math></sub>	116	93.32	94.21	0.9105
EFP <sub><math>\sigma</math></sub>	116	94.50	96.50	0.9281
EFP <sub>a</sub>	116	86.82	93.05	0.8291
EFP <sub><math>\mu+\sigma</math></sub>	232	96.51	97.53	0.9541
EFP <sub><math>\sigma+a</math></sub>	232	95.42	97.81	0.9403
EFP <sub><math>\mu+\sigma+a</math></sub>	348	<b>96.69</b>	<b>98.22</b>	<b>0.9601</b>
<i>Feature profiles with attribute 'moment of inertia'</i>				
EFP <sub><math>\mu</math></sub>	36	83.82	88.71	0.792
EFP <sub><math>\sigma</math></sub>	36	85.20	89.80	0.8111
EFP <sub>a</sub>	36	78.21	86.65	0.7249
EFP <sub><math>\mu+\sigma</math></sub>	72	87.70	90.42	0.8407
EFP <sub><math>\sigma+a</math></sub>	72	84.76	89.17	0.8038
EFP <sub><math>\mu+\sigma+a</math></sub>	108	<b>88.21</b>	<b>90.59</b>	<b>0.8470</b>
<i>Feature profiles with 'area' + 'moment of inertia'</i>				
EFP <sub><math>\mu</math></sub>	152	94.22	94.61	0.9222
EFP <sub><math>\sigma</math></sub>	152	94.63	97.40	0.9303
EFP <sub>a</sub>	152	85.11	92.75	0.8083
EFP <sub><math>\mu+\sigma</math></sub>	304	97.06	98.22	0.9613
EFP <sub><math>\sigma+a</math></sub>	304	96.38	98.10	0.9526
EFP <sub><math>\mu+\sigma+a</math></sub>	456	<b>97.01</b>	<b>98.34</b>	<b>0.9607</b>

of using other feature extraction techniques instead of the above PCA, as mentioned at the end of Section III-B. In details, we exploited the non-linear KPCA technique using RBF kernel with a bandwidth estimated from the mean pairwise distances from 1000 random pixels. Then, 24 KPCA components (involving 99% of total variance) were preserved for the generation of EAP and EFP features. Our results proved that FPs again outperformed APs in terms of classification performance (generally improve about 3% in  $\kappa$  coefficient). Hence, the effectiveness of FPs when combined with different feature selection techniques for spatial-spectral hyperspectral image classification is validated.

#### D. Complexity analysis

From a practical perspective, the only difference between AP construction and FP construction is that the former outputs pixel values, while the latter outputs attribute/feature values. As such, in case a single and the same attribute is used both for filtering and description, their computational complexity and feature vector lengths are identical. Nevertheless, FP technique possesses a much higher discrimination capacity, since given an image filtered w.r.t. an attribute  $t$ , AP only outputs the resulting pixel gray values that are tightly dependent on the chosen attribute. FP on the other hand, can output the values of the attribute  $t$  per pixel, as well as any other arbitrarily selected features; e.g. given an image filtered w.r.t. area, it can output the area, mean, standard deviation, moment of inertia per pixel. Naturally, if one selects to output  $n$  properties (features) per CC, this will lead to an additional computational cost w.r.t. AP, but only for the amount required to calculate  $n - 1$  additional

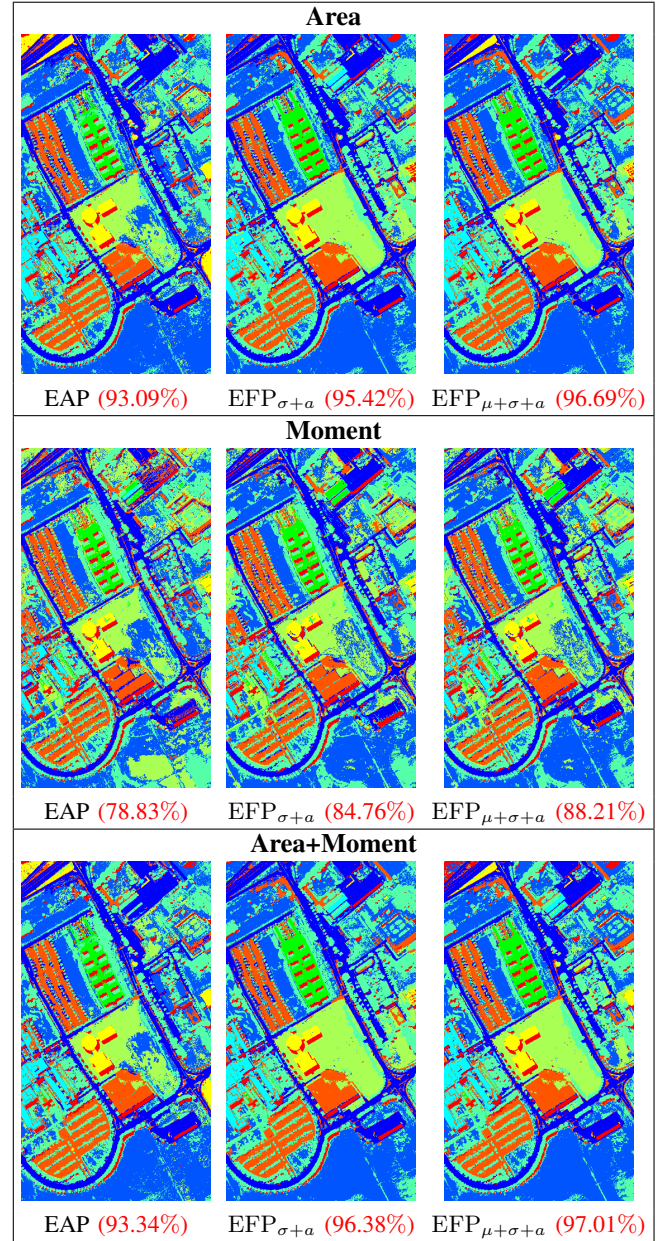


Fig. 5. Classification maps of Pavia University data obtained by different EAP and EFP configurations. Each red number is OA produced by the related method.

attributes per node, which is practically insignificant since the main bottleneck of AP construction is the tree formation. Moreover, longer feature vector will of course bring a further computational cost as far as classification is concerned. To support these remarks, Table V shows the amount of computational time for feature extraction and classification required by different AP and EP techniques from our experiments conducted on the Pavia University data.

#### V. CONCLUSION

The focus of this paper has been on APs, a powerful morphological spectral-spatial pixel description tool, and especially on how to improve their content description performance and flexibility without sacrificing efficiency.

TABLE III  
PER-CLASS CLASSIFICATION PERFORMANCE OF REYKJAVIK DATA YIELDED BY AP AND FP METHODS.

Class	Area			Moment			Area + Moment		
	AP	FP	(+/-)	AP	FP	(+/-)	AP	FP	(+/-)
Residential	70.32	73.01	<b>+2.69</b>	57.80	70.34	<b>+12.54</b>	73.52	73.01	-0.51
Soil	82.60	93.77	<b>+11.17</b>	74.46	84.58	<b>+10.12</b>	83.47	95.04	<b>+11.57</b>
Shadow	83.93	86.19	<b>+2.26</b>	83.83	82.96	-0.87	82.69	83.60	<b>+0.91</b>
Commercial	93.90	93.97	<b>+0.07</b>	83.14	87.95	<b>+4.81</b>	94.00	94.00	-
Highway	49.31	55.66	<b>+6.35</b>	44.53	51.04	<b>+6.51</b>	51.92	54.58	<b>+2.66</b>
Road	84.41	83.43	-0.98	67.23	80.85	<b>+13.62</b>	87.22	87.26	<b>+0.04</b>
OA	81.53	85.81	<b>+4.28</b>	72.51	79.98	<b>+7.47</b>	82.54	86.23	<b>+3.69</b>
AA	77.41	81.01	<b>+3.60</b>	68.50	76.29	<b>+7.79</b>	78.80	81.25	<b>+2.45</b>
$\kappa \times 100$	76.66	81.91	<b>+5.25</b>	65.55	74.69	<b>+9.14</b>	77.93	82.46	<b>+4.53</b>

TABLE IV  
PER-CLASS CLASSIFICATION PERFORMANCE OF PAVIA UNIVERSITY DATA YIELDED BY AP AND FP METHODS.

Class	Area			Moment			Area + Moment		
	EAP	EFP	(+/-)	EAP	EFP	(+/-)	EAP	EFP	(+/-)
Asphalt	95.19	95.36	<b>+0.17</b>	92.43	94.37	<b>+1.94</b>	95.37	95.73	<b>+0.36</b>
Meadows	94.98	95.67	<b>+0.69</b>	67.27	85.27	<b>+18.00</b>	95.02	95.59	<b>+0.57</b>
Gravel	92.62	98.05	<b>+5.43</b>	70.89	59.60	-11.29	88.14	98.09	<b>+9.95</b>
Tree	97.00	97.98	<b>+0.98</b>	99.35	99.48	<b>+0.13</b>	99.25	97.88	-1.37
Metal sheets	99.03	99.85	<b>+0.82</b>	99.48	99.85	<b>+0.37</b>	99.63	99.70	<b>+0.07</b>
Bare soil	72.58	99.13	<b>+26.55</b>	65.30	80.29	<b>+14.99</b>	74.03	99.18	<b>+25.15</b>
Bitumen	99.92	99.92	-	99.77	98.35	-1.42	100	99.92	-0.08
Bricks	98.75	99.43	<b>+0.68</b>	98.32	99.05	<b>+0.73</b>	99.59	99.46	-0.13
Shadows	98.63	98.63	-	99.58	99.05	-0.53	98.52	99.47	<b>+0.95</b>
OA	93.09	96.96	<b>+3.87</b>	78.83	88.21	<b>+9.38</b>	93.34	97.01	<b>+3.67</b>
AA	94.30	98.22	<b>+3.92</b>	88.04	90.59	<b>+2.55</b>	94.39	98.34	<b>+3.95</b>
$\kappa \times 100$	90.84	96.01	<b>+5.07</b>	73.22	84.70	<b>+11.48</b>	91.16	96.07	<b>+4.91</b>

TABLE V  
FEATURE EXTRACTION AND CLASSIFICATION TIME OF DIFFERENT AP AND FP TECHNIQUES APPLIED TO THE PAVIA UNIVERSITY DATA.

Method	Dimension	Execution time (s)	
		Feature extraction	Classification
Attribute profiles			
EAP <sub>a</sub>	116	17.1	4.8
EAP <sub>i</sub>	36	7.8	3.2
EAP <sub>a+i</sub>	152	20.5	5.4
Feature profiles with attribute ‘area’			
EFP <sub>a</sub>	116	17.1	4.8
EFP <sub>μ+σ+a</sub>	348	17.7	7.7
Feature profiles with attribute ‘moment of inertia’			
EFP <sub>a</sub>	36	7.9	3.2
EFP <sub>μ+σ+a</sub>	108	8.6	4.3
Feature profiles with ‘area’ + ‘moment of inertia’			
EFP <sub>a</sub>	152	20.4	5.3
EFP <sub>μ+σ+a</sub>	456	21.6	9.3

APs describe pixels through the values they obtain after connected filtering w.r.t. arbitrary attributes. The resulting pixel values depend on the properties of the CC containing them, and as such possess a significant descriptive power. We propose to further elevate this power, by describing every pixel not through its gray value but through the actual attribute values characterizing the CC that contains it.

This approach enables us not only to capture the finer nuances between different regions/objects, but also to extract multiple arbitrary properties (features) per CC, as opposed to the single gray level in the case of APs; hence leading to a much more flexible solution. The proposed approach has

been validated using one panchromatic and one hyperspectral data set. It has been compared against the standard AP and outperformed not only it but the recently introduced extinction profile as well. Moreover, since the only computational difference between AP and FP techniques involves the type of output value, i.e. attribute value instead of pixel value, this performance improvement entails no significant additional computational cost either.

Future work will focus on exploring alternative tree representations as well as the combination of FPs with deep neural networks.

## VI. ACKNOWLEDGMENT

This work was supported by the French Agence Nationale de la Recherche (ANR) under the reference ANR-13-JS02-0005-01 (Asterix project) and the Région Bretagne grant. It was also supported by the BAGEP Award of the Science Academy and the Tubitak grant 115E857. The authors would like to thank Prof. Jon Atli Benediktsson and Prof. Paolo Gamba for making available the Reykjavik image and the hyperspectral Pavia University data.

## REFERENCES

- [1] L. Bruzzone and B. Demir, "A review of modern approaches to classification of remote sensing data," in *Land Use and Land Cover Mapping in Europe*. Springer, 2014, pp. 127–143.
- [2] M. Dalla Mura, J. A. Benediktsson, N. Waske, and L. Bruzzone, "Morphological attribute profiles for the analysis of very high resolution images," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 10, pp. 3747–3762, 2010.



- [3] M. Dalla Mura, J. A. Benediktsson, B. Waske, and L. Bruzzone, "Extended profiles with morphological attribute filters for the analysis of hyperspectral data," *Int. J. Remote Sens.*, vol. 31, no. 22, pp. 5975–5991, 2010.
- [4] M. Dalla Mura, J. A. Benediktsson, and L. Bruzzone, "Self-dual attribute profiles for the analysis of remote sensing images," in *Int. Symp. Math. Morph. Appl. Sig. Image Proc.* Springer, 2011, pp. 320–330.
- [5] G. Cavallaro, M. Dalla Mura, J. A. Benediktsson, and A. Plaza, "Remote sensing image classification using attribute filters defined over the tree of shapes," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 7, pp. 3899–3911, 2016.
- [6] P. Bosilj, B. B. Damodaran, E. Aptoula, M. Dalla Mura, and S. Lefèvre, "Attribute profiles from partitioning trees," in *Int. Symp. Math. Morph. and Its Appl. to Sig. Image Proc.* Springer, 2017, pp. 381–392.
- [7] B. Demir and L. Bruzzone, "Histogram-based attribute profiles for classification of very high resolution remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2096–2107, 2016.
- [8] M.-T. Pham, S. Lefèvre, E. Aptoula, and B. B. Damodaran, "Classification of VHR remote sensing images using local feature-based attribute profiles," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*. IEEE, 2017.
- [9] M.-T. Pham, S. Lefèvre, and E. Aptoula, "Local feature-based attribute profiles for optical remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, 2017. [Online]. Available: <http://dx.doi.org/10.1109/TGRS.2017.2761402>
- [10] E. Aptoula, "Hyperspectral image classification with multidimensional attribute profiles," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 10, pp. 2031–2035, 2015.
- [11] B. Song, J. Li, M. Dalla Mura, P. Li, A. Plaza, J. M. Bioucas-Dias, J. A. Benediktsson, and J. Chanussot, "Remotely sensed image classification using sparse representations of morphological attribute profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 8, pp. 5122–5136, 2014.
- [12] E. Aptoula, M. C. Ozdemir, and B. Yanikoglu, "Deep learning with attribute profiles for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 12, pp. 1970–1974, 2016.
- [13] P. Ghamisi, R. Souza, J. A. Benediktsson, X. X. Zhu, L. Rittner, and R. A. Lotufo, "Extinction profiles for the classification of remote sensing data," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 10, pp. 5631–5645, 2016.
- [14] E. Aptoula, M. Dalla Mura, and S. Lefèvre, "Vector attribute profiles for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 6, pp. 3208–3220, 2016.
- [15] P. R. Marpu, M. Pedergnana, M. D. Mura, S. Peeters, J. A. Benediktsson, and L. Bruzzone, "Classification of hyperspectral data using extended attribute profiles based on supervised and unsupervised feature extraction techniques," *Int. J. Image Data Fusion*, vol. 3, no. 3, pp. 269–298, 2012.
- [16] P. Ghamisi, J. A. Benediktsson, and J. R. Sveinsson, "Automatic spectral-spatial classification framework based on attribute profiles and supervised feature extraction," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 9, pp. 5771–5782, 2014.
- [17] S. Bernabe, P. R. Marpu, A. Plaza, M. Dalla Mura, and J. A. Benediktsson, "Spectral-spatial classification of multispectral images using kernel feature space representation," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 1, pp. 288–292, 2014.
- [18] D. Lungu, S. Prasad, M. M. Crawford, and O. Ersoy, "Manifold-learning-based feature extraction for classification of hyperspectral data: A review of advances in manifold learning," *IEEE Signal Proc. Mag.*, vol. 31, no. 1, pp. 55–66, 2014.
- [19] X. Kang, S. Li, and J. A. Benediktsson, "Feature extraction of hyperspectral images with image fusion and recursive filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 6, pp. 3742–3752, 2014.
- [20] X. Kang, S. Li, L. Fang, and J. A. Benediktsson, "Intrinsic image decomposition for feature extraction of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 2241–2253, 2015.
- [21] J. Xia, L. Bombrun, T. Adali, Y. Berthoumieu, and C. Germain, "Spectral-spatial classification of hyperspectral images using ICA and edge-preserving filter via an ensemble strategy," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4971–4982, 2016.
- [22] B. Sun, X. Kang, S. Li, and J. A. Benediktsson, "Random-walker-based collaborative learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 1, pp. 212–222, 2017.
- [23] P. Ghamisi, M. Dalla Mura, and J. A. Benediktsson, "A survey on spectral-spatial classification techniques based on attribute profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2335–2353, 2015.
- [24] D. ping Tian, "A review on image feature extraction and representation techniques," *Int. J. Mul. Ub. Engin.*, vol. 8, no. 4, pp. 385–396, 2013.
- [25] J. Tang, S. Alelyani, and H. Liu, "Feature selection for classification: A review," *Data Classification: Algorithms and Applications*, p. 37, 2014.
- [26] I. Jolliffe, *Principal component analysis*. Wiley Online Library, 2002.
- [27] M. Dalla Mura, A. Villa, J. A. Benediktsson, J. Chanussot, and L. Bruzzone, "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis," *IEEE Geosci. Remote Sens. Lett.*, vol. 8, no. 3, pp. 542–546, 2011.
- [28] G. Cavallaro, M. Dalla Mura, J. A. Benediktsson, and L. Bruzzone, "Extended self-dual attribute profiles for the classification of hyperspectral images," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 8, pp. 1690–1694, 2015.
- [29] A. Liaw and M. Wiener, "Classification and regression by randomForest," *R news*, vol. 2, no. 3, pp. 18–22, 2002.



**Minh-Tan Pham** (S'13-M'17) received the M.Eng. and M.Res. degrees in electronics and telecommunications from the Institute Mines-Telecom, Telecom Bretagne, Brest, France in 2013. He obtained his Ph.D. in Information and Image processing from Telecom Bretagne in collaboration with the French Space Agency (CNES) in 2016. He is now a post-doctoral researcher at the OBELIX team, IRISA laboratory, Vannes, France. His research interests include image processing, computer vision and machine learning applied to remote sensing imagery

with the current focus on texture analysis, mathematical morphology, hierarchical representation and deep networks for feature extraction, object detection and classification of remote sensing data. He actually serves as reviewer of IEEE Transactions on Geosciences and Remote Sensing, IEEE Geosciences and Remote Sensing Letters and several MDPI journals including Remote Sensing, Sensors and Electronics.



**Erchan Aptoula** received the B.Sc. degree in computer engineering from Galatasaray University, Turkey in 2004 and the M.Sc. and Ph.D. degrees in computer science from Strasbourg University, France in 2005 and 2008 respectively. He is currently an associate professor at the Institute of Information Technologies of Gebze Technical University at Turkey, working on mathematical morphology, hyper-spectral image analysis as well as on content-based image description and retrieval.



**Sébastien Lefèvre** graduated from University of Technology of Compiègne (M.Sc. and Eng. degrees, 1999), University of Tours (PhD, 2002) and University of Strasbourg (HDR, 2009). From 2003 to 2010, he was an Associate Professor in the Department of Computer Sciences and the Image Sciences, Computer Sciences and Remote Sensing Laboratory (LSIIT), University of Strasbourg-CNRS. In 2009–2010, he was an INRIA invited scientist within the TEXMEX team of IRISA/INRIA Rennes. In 2010, he joined the University Bretagne Sud as a Full

Professor in Computer Science, in the Institute of Technology of Vannes and the Institute for Research in Computer Science and Random Systems (IRISA). Within IRISA, he is leading the OBELIX team dedicated to image analysis and machine learning for remote sensing and earth observation ([www.irisa.fr/obelix](http://www.irisa.fr/obelix)). He has coauthored more than 110 papers in image analysis and pattern recognition. His current research interests are mainly related to hierarchical image analysis and deep learning applied to remote sensing of environment. He was co-chair of GEOBIA 2016 and is co-chair of JURSE 2019.